DOCUMENT RESUME

ED 251 827 CS 007 887

TITLE Testing and Evaluation in Reading and Communication Skills: Abstracts of Doctoral Dissertations Published in "Dissertation Abstracts International," July through December 1984 (Vol. 45 Nos. 1 through 6).

INSTITUTION ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill.

PUB DATE 84

NOTE 12p.; Pages may be marginally legible.

PUB TYPE Reference Materials - Bibliographies (131)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Academic Achievement; Annotated Bibliographies; Doctoral Dissertations; Educational Assessment; *Educational Research; Elementary Secondary Education; English Instruction; Higher Education; Minimum Competency Testing; Miscue Analysis; *Program Evaluation; Reading Comprehension; Reading Research; Reading Tests; State Programs; Teacher Evaluation; *Test Construction; *Testing Problems; Test Reliability; *Test Validity; *Writing Evaluation

ABSTRACT

This collection of abstracts is part of a continuing series providing information on recent doctoral dissertations. The 19 titles deal with a variety of topics, including the following: (1) the effectiveness of English placement examinations used at five junior colleges in California; (2) the direction and distance of context required by college undergraduate students to complete cloze cohesive reference items in passages presented by computer; (3) the influence of rater characteristics on composition evaluation practices; (4) the development of a test of content area reading; (5) the prediction of freshman composition grades at a community college; (6) the effectiveness of the Reading Miscue Inventory and the Reading Appraisal Guide in graduate reading programs; (7) the performance of above average and below average readers on a program-dependent mastery test; (8) the measurement of writing abilities in grades seven, eight, and nine; (9) speech act taxonomy as a tool for ethnographic description; (10) the development and evaluation of an instrument to assess sex stereotyping at the primary reading level, and its application to basal reading textbooks used in the Philadelphia (Pennsylvania) schools; and (11) statewide minimum competency tests in writing skills. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED251827

Testing and Evaluation in Reading and Communication Skills:

Abstracts of Doctoral Dissertations Published in Dissertation Abstracts International, July through December 1984 (Vol. 45 Nos. 1 through 6).

Compiled by the Staff
of the
ERIC Clearinghouse on Reading and Communication Skills

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

UMI

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 007 887

The dissertation titles contained here are published with permission of the University Microfilms International, publishers of Dissertation Abstracts International (copyright © 1984 by University Microfilms International) and may not be reproduced without their proper permission.

This bibliography has been compiled as part of a continuing series designed to make information on relevant dissertations available to users of the ERIC system. Monthly issues of Dissertation Abstracts International are reviewed in order to compile abstracts of dissertations on related topics, which thus become accessible in searches of the ERIC data base. Ordering information for dissertations themselves is included at the end of the bibliography.

Abstracts of the following dissertations are included in this collection:

Chase, Dorothy Doreen
AN ENGLISH PLACEMENT EXAMINATION: AN ANALYSIS OF THE EFFECTIVENESS OF THE EXAMINATION AS USED AT FIVE JUNIOR COLLEGES IN CALIFORNIA

DeGrace, William Gail
READING PROFICIENCY DISCRIMINATORS DERIVED FROM DEFINITIONS OF READING BASED ON A NATIONAL SURVEY OF COMPETENCY TESTS FROM ONE HUNDRED THIRTY-THREE LARGE URBAN SCHOOL DISTRICTS KINDERGARTEN

Ewell, Robert Norris
THE EFFECTIVENESS OF ENGLISH PROFICIENCY MEASURES IN PREDICTED ACADEMIC PERFORMANCE OF INTERNATIONAL MILITARY OFFICERS AT A U.S. AIR FORCE PROFESSIONAL MILITARY EDUCATION SCHOOL

Fortier, Gilles Edouard
DIRECTION AND DISTANCE OF CONTEXT REQUIRED BY COLLEGE UNDERGRADUATES TO COMPLETE CLOZE COHESIVE REFERENCE ITEMS IN PASSAGES PRESENTED BY MICROCOMPUTER

Helgren, Valerie Ann
AN ANALYSIS OF ALTERNATE FORM RELIABILITY OF THREE COMMERCIALLY PREPARED INFORMAL READING INVENTORIES

Hrach, Elaine
THE INFLUENCE OF RATER CHARACTERISTICS ON COMPOSITION EVALUATION PRACTICES

Johnson, Juanita Jamison
EVALUATION OF THE READING PROGRAM OF THE GUILFORD COUNTY SCHOOL SYSTEM

Kahn, Harold Richard
UTILIZATION OF STATEWIDE ASSESSMENT PROGRAM READING RESULTS IN SELECT CALIFORNIA ELEMENTARY SCHOOLS

Knickerbocker, Joan Lee
THE DEVELOPMENT OF A TEST OF CONTENT-AREA READING

Lamb, Bill Henry
THE PREDICTION OF FRESHMAN COMPOSITION GRADES AT A COMMUNITY COLLEGE: A CORRELATIONAL STUDY BASED ON A NONCOMPUTATIONAL READABILITY SCALE

Long, Patricia Catherine
THE EFFECTIVENESS OF THE READING MISCUE INVENTORY AND THE READING APPRAISAL GUIDE IN GRADUATE READING PROGRAMS

Lyons, Kevin Michael
A VALIDITY STUDY OF A DOMAIN-REFERENCED MINIMUM COMPETENCY READING TEST FOR SECONDARY LEVEL STUDENTS AND A PROPOSED MODEL FOR ASSESSING THE CONTEXT DEPENDENCE, RELIABILITY, CONTENT AND CONSTRUCT VALIDITY OF SIMILAR CRITERION-REFERENCED READING MEASURES

McCrea, Darla V.
AN ANALYSIS OF THE PERFORMANCE OF ABOVE AVERAGE READERS AND BELOW AVERAGE READERS ON A PROGRAM-DEPENDEN. MASTERY TEST

McQuillan, Mark Kellogg
THE MEASUREMENT OF WRITING ABILITIES IN GRADES 7, 8, AND 9

Reiss, Nira
SPEECH ACT TAXONOMY AS A TOOL FOR ETHNOGRAPHIC DESCRIPTION: AN ANALYSIS BASED ON VIDEOTAPES OF CONTINUOUS BEHAVIOR IN TWO NEW YORK HOUSEHOLDS

Rudnick, Charlotte Ruth
THE DEVELOPMENT AND EVALUATION OF AN INSTRUMENT TO ASSESS SEX STEREOTYPING AT THE PRIMARY READING LEVEL AND THE ANALYSIS OF BASAL READING TEXTBOOKS USED IN THE PHILADELPHIA SCHOOLS COMPARING THE EARLY EDITIONS PUBLISHED FROM 1965-1971 WITH LATER EDITIONS PUBLISHED FROM 1979-1983

Savale, Zoila Ann
STATEWIDE MINIMUM COMPETENCY TESTS IN WRITING SKILLS: HOW COMPETENT ARE THE TESTS?

Sonnenblick, Carol Anne
READABILITY AND COMPREHENSION: VARIABLES IN TEXT ANALYSIS

Telfer, Richard J.
THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST OF TEACHERS' KNOWLEDGE OF THE TEACHING OF READING COMPREHENSION

AN ENGLISH PLACEMENT EXAMINATION: AN ANALYSIS OF THE EFFECTIVENESS OF THE EXAMINATION AS USED AT FIVE JUNIOR COLLEGES IN CALIFORNIA

Order No. DA8412452

CHASE, DOROTHY DOREEN, PH.D. *Brigham Young University*, 1984. 159pp. Chairman: Milford C. Cottrell

The purpose of this study was to develop means, standard deviations and percentile norms for an English Placement Examination. The examination consisted of two parts, an objective examination and a non-objective examination. Five junior colleges and their incoming students were involved in the norming of the examination.

The examination was found to have predictive value, between student's raw scores on the English Placement Examination and teacher assigned grades, at two colleges. At the third college there was sufficient correlation to merit further use and analysis of the test.

The results obtained at the fourth college showed little correlation. However, when the course levels and raw examination scores were analyzed, for each individual college, a high correlation was found.

READING PROFICIENCY DISCRIMINATORS DERIVED FROM DEFINITIONS OF READING BASED ON A NATIONAL SURVEY OF COMPETENCY TESTS FROM ONE HUNDRED THIRTY-THREE LARGE URBAN SCHOOL DISTRICTS KINDERGARTEN

Order No. DA8411011

DEGRACE, WILLIAM GAIL, ED.D. *University of San Francisco*, 1983. 324pp. Chairperson: Dr. Joan Hyman

This is one of four collaborative multi-phased studies which investigated the national status of competency-based reading instruction at four grade levels: kindergarten (DeGrace, 1983), grade three (Smith, 1983), grade six (Baker, 1983), and grade eleven (Eggers, 1983). This particular study focused on kindergarten competencies in reading comprehension.

Each of the four studies was divided into three substudies. The results of the first substudy, a survey of the status of competency-based reading comprehension instruction in the one hundred thirty-three largest school districts in the nation, indicated how the competencies were mandated, derived, used, and evaluated by each of these districts. In the second substudy the competencies and competency test items were evaluated by panels of internationally recognized reading experts. The third substudy was an in depth analysis of actual competency test items submitted by the surveyed school districts. To accomplish this analysis an instrument was developed and administered to students at each selected grade level. Three types of comprehension test items selected from the school districts' submissions and identified by the reading experts as being (1) most relevant, (2) least relevant, and (3) items on which there was no agreement by the experts, were included.

The findings of this study identified individual comprehension items as well as clusters of items that discriminated proficient from less proficient readers at the kindergarten level. Agreement across experts' evaluations of reading comprehension test items and the validation of those items through student testing resulted in item clusters and item discriminability and yielded a core of reading comprehension items. These items could form the basis of a national data bank of reading comprehension test items and therefore "true" reading comprehension instructional objectives. In other words these items indicate what a cross section of reading experts, students, and school districts used as an operational definition of reading comprehension in American public schools.

THE EFFECTIVENESS OF ENGLISH PROFICIENCY MEASURES IN PREDICTING ACADEMIC PERFORMANCE OF INTERNATIONAL MILITARY OFFICERS AT A U.S. AIR FORCE PROFESSIONAL MILITARY EDUCATION SCHOOL

Order No. DA8415499

EWELL, ROBERT NORRIS, ED.D. *Auburn University*, 1984. 98pp. Director: Wilbur A. Tincher

The purpose of the study was to determine how much of the variation in academic performance of international officers at the U.S. Air Force's Squadron Officer School (SOS) was attributable to their scores on English proficiency examinations. Subpurposes involved investigating the relative effectiveness of the English examinations.

Fifty-one international officers of two successive SOS classes in March and May, 1983, were administered four English proficiency tests: the English Comprehension Level (ECL) test, which is used by the Department of Defense to screen prospective students, the International Officer School (IOS) Aural test, the IOS Oral test, and the IOS Reading test. Student scores on those tests were compared with their SOS final test average. Data were analyzed by means of multiple regression techniques with R^2 values and F-ratios the key points of interest.

The four English tests together accounted for 16 percent of the variance in final test average. The Reading test was the single best predictor at 13 percent; the ECL alone accounted for only 6 percent. Among tests and combinations of tests that were statistically significant predictors of academic performance at the .05 level were Reading alone and Reading and ECL. It was noted that the predictors were moderately intercorrelated (.66-.80) and that there were some suppression effects.

It was concluded that while English proficiency may be a necessary condition for academic performance, differences in scores on proficiency examinations account for a relatively small proportion of the differences in academic test scores. It also was concluded that the Reading test was the best predictor because it most accurately simulated the academic and testing environment of SOS. It was recommended that institutions not rely on English tests as sole admission criteria for international students. It was also recommended that language skills essential to specific curricula be identified, tested, and remediated, and that poorly performing international students be counseled to determine whether the cause of poor performance is language or some other factor.

DIRECTION AND DISTANCE OF CONTEXT REQUIRED BY COLLEGE UNDERGRADUATES TO COMPLETE CLOZE COHESIVE REFERENCE ITEMS IN PASSAGES PRESENTED BY MICROCOMPUTER

Order No. DA8421109

FORTIER, GILLES EDOUARD, PH.D. *University of Georgia*, 1984. 121pp. Director: George E. Mason

The purpose of this study was to determine whether direction and distance of context had an effect on the way readers process text information before attempting to complete the blanks in a cloze test. Based on the Halliday and Hasan (1976) cohesion theory, 25 cloze tasks were constructed in which the reference item of a pair of cohesive ties was deleted. In 20 cases, the presupposed item was anaphoric, and in five cases, it was cataphoric. The cloze tasks were taken from an expository text found to be at the 10th grade readability level.

A microcomputer program was written by the investigator. It displayed the anaphoric or cataphoric context word-by-word, but only upon request by the subjects, who were 38 volunteer undergraduate students enrolled at the University of Georgia. The reading ability of

the subjects was measured with the *Nelson-Denny Reading Test*. Their percentile rank was used to divide them into two reading groups, the less proficient readers and the more proficient readers.

The results indicated that more proficient readers did not complete successfully a greater number of blanks than less proficient readers. Nor did they ask for significantly more words before completing the blanks. Less proficient readers reached the presupposed items significantly more often and a moderately low negative correlation was found between the number of times they reached the presupposed items and the number of blanks they completed successfully, but no such significant correlation was found for more proficient readers. The direction of context did not affect the number of blanks completed successfully by either type. But the direction of context affected both the number of times the presupposed items were reached by the less proficient readers and the number of words they asked before completing the blanks. A moderate negative correlation was found between the number of times they reached the presupposed items and the number of blanks they completed successfully with anaphoric context, but not with cataphoric. With more proficient readers no significant correlation was found.

It was concluded that regardless of their levels of reading proficiency, (1) college undergraduates process essentially the same information before attempting to complete the blanks in a cloze test, and (2) they process the information differently when the direction of context is anaphoric than when it is cataphoric.

AN ANALYSIS OF ALTERNATE FORM RELIABILITY OF THREE COMMERCIALLY PREPARED INFORMAL READING INVENTORIES

Order No. DA8416328

HELGREN, VALERIE ANN, ED.D. *University of Miami*, 1984. 45pp. Supervisor: Professor Charles T. Mangrum, II

An analysis of alternate form reliability of three commercially prepared informal reading inventories was performed in this study. Seventy-five fourth grade students from two elementary schools were randomly assigned to one of three IRIs and administered Forms A and B of an inventory. This resulted in the allocation of twenty-five students to each IRI. The *Analytical*, *Basic*, and *Ekwall* Reading inventories were used.

Pearson and generalizability coefficients were computed for the data. Results indicated that the coefficients ranged between .61 and .78. Estimated variance components from the generalizability analysis revealed that little error could be directly attributed to the forms, as the largest variance component was attributed to the subjects themselves.

While the results of the study did not reveal perfect reliability, critics cannot condemn these IRIs for being unreliable. Reading researchers need to address the question of what an acceptable level of reliability would be for informal reading inventories.

THE INFLUENCE OF RATER CHARACTERISTICS ON COMPOSITION EVALUATION PRACTICES

Order No. DA8410956

HRACH, ELAINE, PH.D. *Hofstra University*, 1983. 223pp. Chairperson: Estelle S. Gellman

The purpose of this study was to further existing knowledge concerning the evaluation of student writing samples. The evaluation practices of 59 junior and senior high school teachers of writing were examined to determine whether or not patterns of rating would emerge that are related to the rater characteristics of tolerance of ambiguity and/or experience with writing skills.

Hypothesis One predicted that rater judgments of writing quality would be multidimensional. Hypothesis Two sought to determine how the dimensionality associated with a particular evaluator's judgments would be related to experience with writing skills and tolerance of ambiguity. It was predicted that evaluators scoring highest on these variables would utilize more dimensions than other respondents in their evaluations and would weight content dimensions most heavily.

No standards of writing were suggested to the participants of this study. Instead, teachers were asked to read 20 papers on the same topic which were written by high school students. These papers were to be sorted so that papers placed in any one category were similar in terms of writing quality and different from papers placed in other categories.

Three-way multidimensional scaling was employed to uncover the structure underlying teachers' sorting data. A four-dimensional solution offered both goodness-of-fit and interpretability. On the basis of characteristics of papers placed at dimensional extremes, the derived dimensions seemed to represent the following writing attributes: Style, Maturity of thought and expression, Organization and Substance.

The 20 compositions also were rated on specific writing traits by 39 additional teachers. The mean bipolar scale ratings obtained from these subjects were regressed over each paper's projections on the four MDS dimensions. These regression analyses substantiated the original dimensional interpretations.

A canonical correlation analysis was performed to determine relationships between teachers' weights on the MDS dimensions and their scores on the variables "Experience with Writing Skills" and "Tolerance of Ambiguity." The canonical variates formed in this analysis did not account for any significant relationship between rater variables and dimension weights. Thus, Hypothesis Two, which suggested that rater judgments would be a function of experience with writing skills and tolerance of ambiguity was not supported.

EVALUATION OF THE READING PROGRAM OF THE GUILFORD COUNTY SCHOOL SYSTEM

Order No. DA8408997

JOHNSON, JUANITA JAMISON, ED.D. *The University of North Carolina at Greensboro*, 1983. 186pp. Director: Dr. Carol K. Tittle

The purpose of this study was to determine the impact of curricular intervention on the reading program of a local school district. The study was based on the following assumptions: that evaluation is part of a total curriculum improvement process; that principals, teachers, parents and students are among the most important audiences for the evaluation report and should therefore be involved in the evaluation study; and that metaevaluation techniques can be used to improve the validity, reliability and objectivity of local evaluation studies.

Selected literature related to major evaluation models, to the evaluation of reading programs, and to metaevaluation was reviewed. Pretest and posttest survey data were analyzed to test hypotheses relative to a change in the description of the program and the satisfaction of principals, teachers, parents, and students with the program. The reading achievement scores of third and sixth graders in 1979 were compared to 1982 scores for third and sixth graders to test an hypothesis related to student's reading achievement. A metaevaluation study was conducted to validate that determined effects were due to the curricular intervention and not to other threats to internal validity.

The results of the study indicated that teacher's knowledge of the school system's curricular guide had increased, the record keeping of reading skills was more effective, use of the system's curricular guide as a source of reading program goals had increased, student achievement in reading had increased, and teachers were more satisfied with the reading program.

Major conclusions drawn from the study were that a variety of evaluation models exist and are being used to evaluate reading programs; the purposes of the evaluation, the information that is required and the audiences to be served help to determine the models and/or combination of evaluation models that are utilized; evaluations at the local level can be used to determine program effects; and metaevaluation techniques can be used to improve the validity, reliability, and objectivity of local evaluation studies.

UTILIZATION OF STATEWIDE ASSESSMENT PROGRAM READING RESULTS IN SELECT CALIFORNIA ELEMENTARY SCHOOLS

Order No. DA8412572

KAHN, HAROLD RICHARD, ED.D. *University of San Francisco*, 1983. 149pp. Chairperson: Laurence A. Bishop

The accountability movement of the 1960s gave rise to the California Assessment Program. At present, the program administers statewide tests in grades 3, 6, and 12 covering the content areas of reading, written expression, spelling, and mathematics. A major purpose of the program is to provide school and district-level information useful for program improvement.

The objective of this study was to investigate how the public elementary school principals in California utilize the state assessment reading results. The study was designed to explore the influences of factors such as school size, location, ethnic composition, and socioeconomic status of students with regard to local utilization of CAP results. The data for this investigation were provided by a cross-sectional time-bound survey of a representative sample of 292 elementary school principals in the state. The questionnaire was divided for purposes of analysis into three rational categories: dissemination, usefulness, and influence of results.

The results of the study show that a majority of the elementary school principals are receptive to CAP. They use the CAP reading results in conjunction with other information to evaluate their programs. There were other uses of CAP results by the principals. For example, among the "innovative" uses of CAP results were (1) use of the results to gain recognition and support for programs, and (2) confirmatory use of the results to reduce uncertainties about local reading programs. A major finding of the study was that school characteristics, such as ethnic make up, location of the school, age, gender, years of experience of the principal, and academic performance of the students were related to perceived usefulness of CAP results.

In general the results of this study show that CAP reading results are utilized by the elementary school principals for multiple purposes and that how the results are disseminated and used at the school sites depends on the characteristics of the schools.

THE DEVELOPMENT OF A TEST OF CONTENT AREA READING

Order No. DA8409971

KNICKERBOCKER, JOAN LEE, PH.D. *University of Illinois at Urbana-Champaign*, 1983. 126pp. Adviser: Alan L. Madsen

The primary purpose of this study was to develop and pilot test an objective instrument designed to assess secondary school subject area teachers' knowledge of reading as it applies to their teaching. A secondary purpose of the study was to determine how subject area taught, grade level taught, length of teaching experience, attendance in college or university courses emphasizing reading instruction, attendance in course work beyond the bachelor's degree and personal judgment of the anticipated benefit from further training in reading, relate to performance level on the test. The instrument is intended to be utilized in pre-service and in-service teacher training.

The first phase of the research included a survey of the conceptual and research literature in secondary level reading to determine those instructional practices which constitute content area reading. Instrument development included writing a content-process outline derived from the identified universe. Based on the outline, a pool of preliminary multiple-choice items were written. The preliminary instrument and content outline were submitted to authorities in the field of content area reading for content validation. Based on their suggestions the instrument was revised. The instrument was pilot tested with 145 volunteer teachers. The instrument was again revised based on the data analysis. The revised 40-item instrument was pilot tested with 266 volunteer secondary school subject area teachers.

The reliability of the instrument was .770 (Kuder-Richardson # 20); standard error of measurement, 2.77; standard deviation, 5.79; mean score, 24.55; and median score, 25.06.

Teachers who had taken a course emphasizing reading instruction performed significantly better than teachers who had not. Teachers of English/Language Arts performed significantly better than teachers of other subject areas. Grade level taught, length of teaching experience, attendance in college or university courses beyond the bachelor's degree and perceived benefit from training in reading were not found to be significantly related to test performance.

The results of the study support the recommendation for training in appropriate reading instruction for all secondary level pre-service and in-service teachers.

THE PREDICTION OF FRESHMAN COMPOSITION GRADES AT A COMMUNITY COLLEGE: A CORRELATIONAL STUDY BASED ON A NONCOMPUTATIONAL READABILITY SCALE

Order No. DA8415654

LAMB, BILL HENRY, PH.D. *Kansas State University*, 1984. 187pp.

The purpose of this study was to test the effectiveness of the Rauding Scale of Prose Difficulty applied to a sample of a freshman students' writings in order to predict their Written Communications I grade. By applying an index to a writing sample of 300 words, it was possible to rate an individual's potential for success. It was possible to identify students in need of English remediation.

The study was comprised of two randomly divided groups with 110 subjects in Group 1 and 114 subjects in Group 2. The method of this study involved six steps. First, students completed a writing sample consisting of approximately 300 words. Three qualified raters identified by the Rauding Scale Qualification Test then rated each sample and applied a grade level index to the writing. The first group was analyzed and at the conclusion of the semester, the mean grade level of each student was correlated to his/her course grade. The correlation ($r = +.467$) was found to be significant beyond the .001 level.

Based on the significant correlation found in the first group, a Linear Regression Analysis was performed which generated a prediction regression line ($Y' = .200X + 1.031$). The second group was then rated by the same raters using the readability scale. Predicted end course grades were derived using the regression formula. The correlation between the predicted grades and the actual end course grades was found to be $r = +.442$. This figure was also significant beyond the .001 level of confidence.

The implications of this finding were that the Rauding Scale of Prose Difficulty applied to a freshman student writing sample was a viable index for predicting success or failure in Written Communications I at a selected community college.

THE EFFECTIVENESS OF THE READING MISCUE INVENTORY AND THE READING APPRAISAL GUIDE IN GRADUATE READING PROGRAMS

Order No. DA8421975

LONG, PATRICIA CATHERINE, PH.D. *The University of Arizona*, 1984. 206pp. Director: Yetta M. Goodman

The purpose of this study was to examine differences in the effectiveness of two graduate teacher education programs in reading assessment, one group using the Reading Miscue Inventory and the other using one of its simplified forms, the Reading Appraisal Guide.

The main question that is answered in this study is whether it is more effective for teachers to be given training in the Reading Miscue Inventory, or is training in the Reading Appraisal Guide sufficient to enable teachers to carry out competent assessments of children's reading ability?

In the six months of the study's duration, different types of data were collected. These consisted of assessments of children's taped readings of a story by two groups of teachers before (the pretest) and after (the posttest) their respective training programs; anecdotal records of the teachers' views of the programs and the assessment instruments they were using, and observations of the teachers' reading assessments of children selected by them for their practicum.

Quantitative analyses of the pretest and posttest were made; these were based on criteria drawn from the Reading Miscue Inventory manual and the investigator's own miscue analysis of the children's taped readings. They showed that the teachers trained in miscue analysis, as reflected in the Reading Miscue Inventory, were able to

make significantly better assessments of children's reading ability than the teachers trained in the Reading Appraisal Guide.

In addition to the quantitative analysis, written and oral statements made by the teachers during the pretest, posttest and training programs were subjected to qualitative analysis and comparisons. These indicated that both groups' programs had strengthened the teachers' adherence to the Goodman model of reading, but those trained in the use of the Reading Miscue Inventory developed more effective assessment abilities and were more approving of the instruments they used, than were those trained in the use of the Reading Appraisal Guide.

It was concluded that the Reading Miscue Inventory is an appropriate assessment instrument for use in college graduate reading programs. It proved complex and time-consuming to use, but at the same time it enabled teachers to make more accurate, in-depth assessments of children's reading than did the Reading Appraisal Guide. The latter was found to have some serious drawbacks, mostly arising from attempts to make it quicker and easier to use.

A VALIDITY STUDY OF A DOMAIN-REFERENCED MINIMUM COMPETENCY READING TEST FOR SECONDARY LEVEL STUDENTS AND A PROPOSED MODEL FOR ASSESSING THE CONTEXT DEPENDENCE, RELIABILITY, CONTENT AND CONSTRUCT VALIDITY OF SIMILAR CRITERION-REFERENCED READING MEASURES Order No. DA8414711

LYONS, KEVIN MICHAEL, ED.D. *Boston University*, 1981. 302pp. Major Professor: Thomas E. Culliton, Jr.

The purpose of this study was to investigate and describe the reliability and validity of one test of minimum competency in reading designed for secondary level students. Of particular interest were the context-dependence of the instrument and the content and construct validity of the measure. Additionally, the intent of the study was to suggest a practical model for the validation of criterion-referenced reading tests and to describe the limitations of these instruments in practical terms of score interpretation and educational decision-making.

A total sample of 221 eighth-grade students was administered the following measures: (1) *Comprehensive Tests of Basic Skills*, Reading; Level 4; (2) *Short Form Test of Academic Aptitude*, Level 4; (3) *Test Sophistication Scale*. The total sample was randomly divided into two groups. One group, called the control group, was administered the Massachusetts *Assessment of Basic Skills, Reading Form 1* as suggested by test directions. The second group, called the experimental group, was administered an experimental form of the same measure where all reading context was removed by the researcher. Treatment of the test data and a review of related literature enabled the researcher to reach the following conclusions: (A) Context-Dependence of the Intument. (1) The experimental group, which took the competency test form with all reading context removed, achieved a group mean percent correct of 53, which was two-thirds the mean percent correct of 79 achieved by the control group in the normal administration of the 55 item test. (2) A majority of the test items was determined to be flawed due to a substantial lack of context-dependence. (3) Item construction flaws, in terms of identified item writing principles designed to limit the influence of specific test-taking skills, contributed to the substantial lack of context-dependence demonstrated by this test for the study population. (4) The use of indices of context-dependence appeared most useful in isolating poorly constructed reading test items. (5) The noted lack of context-dependence suggested that some skill or skills other than those defined in test definitions were assessed. (B) Relationships. (1) The high intercorrelation of scores on the criterion-referenced reading test, norm-referenced reading test, group I.Q. test, and test-wiseness test suggested an undefined general ability contributed substantially to performance on the four measures. . . . (Author's abstract exceeds stipulated maximum length. Discontinued here with permission of author.) UMI

AN ANALYSIS OF THE PERFORMANCE OF ABOVE AVERAGE READERS AND BELOW AVERAGE READERS ON A PROGRAM-DEPENDENT MASTERY TEST

Order No. DA8420347

MCCREA, DARLA V., PH.D. *The University of Florida*, 1983. 111pp. Chairman: William Hedges

This study examined the relationship between pupil performance on a program-dependent mastery test in reading and overall reading ability, school assignment and their interaction. This study was designed to determine whether a competency-based program in reading reduces the usual normal distribution of reading achievement when all pupils are allowed varying amounts of time to master program objectives.

The dependent variable used was the Ginn Level 10 mastery test (copyright, 1976). Student scores on the total test and eight subtests were analyzed. Independent variables were students' normed reading ability, measured by the Metropolitan Achievement Test (copyright, 1978), schools to which students were assigned, and their interaction. Four hundred and nine tests were collected from grades three through six. Subgroups of above average (n = 121) and below average (n = 65) readers were created for the inferential analyses.

A linear regression model was used to test for significant relationships between the dependent and independent variables. A chi-square analysis tested for a significant difference in the proportion of students from each subgroup who attained mastery criterion. Item analyses (p-values, latent-trait difficulty indices, and goodness-of-fit to the one parameter logistic model) were computed and Pearson product moment correlations were computed between nine variables.

Results of the analyses indicated that MAT was a strong predictor of performance on the mastery test (alpha level .01). School assignment was significant for one subtest (Vocabulary I) and one interaction was significant (Decoding I). A significant difference in the proportions of subgroups achieving mastery favored the above average readers.

The research concluded the performance of below average readers was significantly lower than above average readers despite additional time in instruction. Although varying the amount of instructional time was not a sufficient intervention, the interaction observed for one subtest and the significant relationship between school assignment and performance on another suggest that teachers in some schools have devised interventions which minimize student dependence on overall reading ability.

Research is needed to ascertain which factors were successful interventions. Research is also needed to determine how best to evaluate competency-based instructional programs and testing components which accompany them.

THE MEASUREMENT OF WRITING ABILITIES IN GRADES, 7, 8, AND 9 Order No. DA8416933

MCQUILLAN, MARK KELLOGG, ED.D. *Harvard University*, 1984. 375pp.

How stable are students' scores on different composition tests? Will their scores show uniform response patterns to different topics, or will there be widespread interactions between students and the rhetorical conditions embedded in each topic? Lloyd-Jones (1977) has suggested that widespread interactions are probable, given the developmental differences among children, their diverse rhetorical skills, and the nature of writing ability itself.

There is little evidence to substantiate this claim, however. Until the development of holistic scoring in 1962 (Godshalk, Swineford & Coffman, 1966), low reliabilities have historically confounded scholars' ability to analyze how different writing topics affected students' scores. Even now that high reliabilities are possible on essay tests, experimental research on topic and discourse variation remains scattered. Assessment experts (Cooper, Odell & Courts, 1978) have thus urged investigators to consider how such rhetorical variables as (1) discourse purpose, (2) audience, and (3) form affect students' scores.

To this end, twelve writing tasks, representing three types of discourse (Kinneavy, 1971) and two different speaker-audience

conditions (Gibson, 1969) were randomly administered to a stratified sample of 450 7th, 8th, and 9th grade students. Written as case studies (Hilgars, 1982), the topics were administered over eight weeks and then rated for overall quality in holistic scoring sessions.

Quality measures were then aggregated and analyzed to see if fluctuations in scores were systematically related to differences introduced by: (1) discourse type, (2) "formal" and "informal" audience conditions, (3) the age of the writer, (4) the writer's academic achievement, or (5) some combination thereof. ANOVAs suggested that Lloyd-Jones' hypothesis was valid for early adolescent children, that their group profiles were idiosyncratic. Significant within-group findings showed that discourse aim and audience formality differentially affected the students in nine distinct curriculum divisions. Different writing abilities were evident in each of these groups, suggesting that writing as a multifaceted skill which varies from task to task, depending on the rhetorical elements of the topic, the writers' ages, and their literacy achievement.

SPEECH ACT TAXONOMY AS A TOOL FOR ETHNOGRAPHIC DESCRIPTION: AN ANALYSIS BASED ON VIDEOTAPES OF CONTINUOUS BEHAVIOR IN TWO NEW YORK HOUSEHOLDS

Order No. DA8413019

REISS, NIRA, PH.D. *Columbia University*, 1984. 235pp.

The subject of this study is the designing of measures for ethnographic description which includes speech acts in an etic instrumental approach, oriented toward analyzing the functions of communicative events in relation to the behavior stream. The ultimate goal of designing such measures is to place them within a cultural materialist ethnographic context (Harris 1980). For this purpose I extend Searle's speech act taxonomy (1970) to include intended effects as part of the meaning of speech acts and reduce the emic criteria essential for comprehending and identifying them.

The data base for the study is part of a corpus of videotaped continuous behavior over several days in four New York households recorded by Professor Marvin Harris and a team of students in the course of the study of Patterns of Authority in Low-Income Domiciles. I use transcriptions from one day in each of two English-speaking households, including both communicative events and non-verbal contexts and responses. These are studied in a heuristic way to illustrate a critique of emic approaches, and in a systematic way to apply the taxonomy of speech acts over two sections of data included as Appendices.

In this study I show that the revised taxonomy of speech acts is applicable to an empirical corpus, and that it produces a systematic set of behavioral measures which are potentially productive for cross-cultural comparison. I show that, contrary to claims made by Searle, as well as by Labov (1981), comprehension and production of speech acts do not depend on axiomatic processes according to an elaborate system of felicity rules or conditions. Instead, comprehension depends on control of the functions of speech acts as devices for motivating effects. The conceptual contents and structure of speech acts follow from their instrumental functions.

This dissertation, by looking at attitudes toward Civil Rights, Civil Liberties, and Welfare issues, has suggested that status inconsistency indeed has conceptual possibilities not ascertainable by regression analysis. Rather, the Diamond-Additive technique, as formulated by Keith Hope for social mobility studies, offers a tentative measurement clarification model by allowing one to look at the effects of the status variables and status inconsistency simultaneously, a possibility not offered by regression analysis.

Further discussion, in this dissertation, centers around the issue of status inconsistency as an indicator of stress.

THE DEVELOPMENT AND EVALUATION OF AN INSTRUMENT TO ASSESS SEX STEREOTYPING AT THE PRIMARY READING LEVEL AND THE ANALYSIS OF BASAL READING TEXTBOOKS USED IN THE PHILADELPHIA SCHOOLS COMPARING THE EARLY EDITIONS PUBLISHED FROM 1965-1971 WITH LATER EDITIONS PUBLISHED FROM 1979-1983

Order No. DA8410159

RUDNICK, CHARLOTTE RUTH, ED.D. *Temple University*, 1984. 257pp. Major Adviser: Dr. Eugene Abraham

Purpose. This study determined the nature and change in the depiction of male and female stereotype roles in basal reading textbooks from grades one to three. Editions published from 1965-1971 were compared with editions from 1979-1983.

Procedures. Primary level basal textbooks were selected using two criteria: (1) frequency of use in Philadelphia; (2) publication dates--two editions of each series were used. The five selected were: Lippincott, Harcourt Brace Jovanovich, Houghton Mifflin, Ginn, and Scott Foresman.

The method of investigation was content analysis. The "Sex Stereotyping Checklist" was developed consisting of twelve categories from which the primary hypothesis and subordinate hypotheses evolved. Categories included, among others, fear, aggression, passivity, and subordination. A pilot study demonstrated the utility of the Sex Stereotyping Checklist. Initial validity for the Checklist was accomplished using a panel of experts and reliability was established by the inter-rater method. Content analysis results were reported in frequencies and percentages which were tested for statistical significance.

Conclusions. The analysis of the primary hypothesis showed an increase in proportional representation for female characters analyzed from early to later editions. Ginn showed the greatest level of change in the number of characters depicted. Houghton Mifflin's treatment of the sexes was the most even in the later edition. Lippincott and Scott Foresman made progress while Harcourt Brace showed little change.

The analyses of subordinate hypotheses showed statistical significance in terms of more equal representation of the sexes when comparing early and later editions for fear, aggression, leadership, adequacy, and inadequacy. Passivity, affection, tenderness, subordination, and traditional and non-traditional occupational roles did not show statistically significant changes. Ginn demonstrated the most change in the later edition resulting in a more equal representation of the sexes. Houghton Mifflin's treatment of the sexes was the most evenly proportionate in the later edition. Lippincott and Scott Foresman showed change and Harcourt Brace displayed little evidence of modification.

Recommendations. Reading textbooks should be reviewed to determine if the sexes are portrayed evenly. Publishers should attribute to both sexes the full gamut of human expression.

STATEWIDE MINIMUM COMPETENCY TESTS IN WRITING SKILLS: HOW COMPETENT ARE THE TESTS?

Order No. DA8420012

SAVALE, ZOILA ANN, PH.D. *Indiana University of Pennsylvania*, 1984. 154pp. Chairman: Patrick Hartwell

This dissertation discusses existing minimum competency testing programs, including the concerns of the public, students, teachers, and educational administrators regarding the use of such tests. State-mandated MCT programs are described, and a distinction is drawn between statewide writing MCTs and other statewide writing assessments. The National Assessment of Educational Progress is discussed as an example of the difference between objective referenced and norm referenced testing, and as a potential alternative source of the data currently being collected by MCTs and statewide assessment tests.

Turning specifically to MCTs in writing skills, in Chapter Three a brief discussion of district-administered versus state-administered

writing tests is presented. In Chapter Four, general issues in the testing of writing skills are discussed, including discrete point versus integrated testing, and the choice of writing topics. Since the three most outstanding statewide MCTs in writing are all writing sample tests, and since these three tests are all scored holistically, the remainder of the dissertation is concerned with those tests and the issues relevant to their use, including the various scoring methods available. Finally, holistic scoring, the evaluation technique currently most often used in conducting large-scale writing sample tests, is discussed, some of its limitations are pointed out, and variations are suggested.

READABILITY AND COMPREHENSION: VARIABLES IN TEXT ANALYSIS Order No. DA8411004

SONNENBLICK, CAROL ANNE, ED.D. *Rutgers University The State U. of New Jersey (New Brunswick)*, 1983. 194pp. Chairperson: Josephine Goldsmith

Problem. This study examined the effects of two sets of readability variables on the comprehensibility of text, those associated with traditional readability research, and those derived from discourse analysis and based on Kintsch's model of comprehension.

Procedure. Forty (40) passages from the reading comprehension test of the Tests of General Educational Development (GED) were analyzed to derive traditional readability variables, measures of vocabulary and syntactic complexity. Next, these passages were parsed into their propositional structure and mapped into coherence graphs, thus allowing the quantification of discourse analytic variables. The dependent measures were the average probabilities of GED candidates responding correctly to the test items for each passage. Data was analyzed using stepwise multiple regression procedures with three sets of independent readability variables, traditional (set 1), discourse derived (set 2) and combined (sets 1 and 2). Each of these sets was analyzed for the total, good and poor reader groups.

Findings. The major findings are: (1) Variables derived from discourse analysis explained a larger proportion of the variance in reader comprehension than did traditional readability factors; (2) All analyses were less successful (not statistically significant) for the poor reader group than for the total or good reader samples, regardless of the variables; (3) When old and new variables were combined, number of inferences, long sentences and a measure of proposition density were the strongest predictors of reader performance in comprehension for all readers, explaining approximately 51% of the variance for the total and good reader groups.

Conclusions. For this sample, out-of-school adults, variables derived from discourse analysis were more highly correlated with text difficulty than were the traditional readability variables. When traditional and discourse derived variables were combined, a larger proportion of the variance in comprehension was explained than when either set was correlated by itself.

Discourse analysis in its present form is such a complicated technique that it is only suitable as a research tool. Readability formulas are useful, but as important, they are usable in the practical sense.

THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST OF TEACHERS' KNOWLEDGE OF THE TEACHING OF READING COMPREHENSION Order No. DA8414260

TELFER, RICHARD J., PH.D. *The University of Wisconsin - Madison*, 1984. 193pp. Supervisor: Professor Kenneth L. Dulin

The purpose of this study was to design and develop a test of teachers' knowledge of the teaching of reading comprehension. Initially, the need for such a test was established and the possible benefits of the test were detailed. Subsequently, the question of what knowledge should be included on the test was examined. Finally, the test development process was discussed.

In the initial phase of test development, a comprehensive outline of knowledge of reading was prepared. This outline was then narrowed to knowledge of the teaching of reading comprehension, including knowledge within five categories: (1) the nature of reading comprehension, (2) the nature of materials, (3) the nature of readers, (4) reading comprehension teaching strategies, and (5) application of knowledge of the teaching of reading comprehension.

The development of the test itself began with the writing of items in each of the categories of the detailed test outline. Once the original items were written, they were reviewed and revised. The revised items were then pilot-tested and again revised when necessary. Finally, 50 items were chosen for the final instrument and the final instrument was field-tested.

In the field test the final instrument was administered to 525 undergraduate and graduate students in reading methods courses at four Wisconsin universities. For purposes of analysis, the scores of each examinee were assigned to one of three teaching-experience groups and to one of three coursework groups. An analysis of variance and follow-up Scheffé comparisons showed that the mean scores of the three teaching experience groups (no experience, one to three years teaching experience, and four or more years teaching experience) were significantly different one from another. Also, an analysis of variance for the three coursework groups (no coursework, one course, and two or more courses) showed that the mean scores of these three groups were significantly different one from another. Finally, significant Pearson product-moment correlations were found both between teaching-experience categories and examinees' scores and between coursework categories and examinees' scores, with examinees' scores increasing as teaching-experience categories and coursework categories rose.

Copies of the Dissertations may be obtained by addressing your request to:

University Microfilms International
Dissertation Copies
Post Office Box 1764
Ann Arbor, Michigan 48106

or by telephoning (toll-free) 1-800-521-3042